



HAL
open science

De novo coding variants in the AGO1 gene cause a neurodevelopmental disorder with intellectual disability.

Audrey Schalk, Margot A Cousin, Nikita R Dsouza, Thomas D Challman,
Karen E Wain, Zoe Powis, Kelly Minks, Aurélien Trimouille, Eulalie
Lasseaux, Didier Lacombe, et al.

► **To cite this version:**

Audrey Schalk, Margot A Cousin, Nikita R Dsouza, Thomas D Challman, Karen E Wain, et al..
De novo coding variants in the AGO1 gene cause a neurodevelopmental disorder with intellectual
disability.. Journal of Medical Genetics, 2021, 10.1136/jmedgenet-2021-107751 . hal-03697474

HAL Id: hal-03697474

<https://hal.science/hal-03697474>

Submitted on 21 Nov 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

De novo coding variants in the *AGO1* gene cause a neurodevelopmental disorder with intellectual disability

Audrey Schalk, MD^{1,*}, Margot A. Cousin, PhD^{2,3,*}, Thomas D. Challman, MD⁴, Karen E. Wain, MS⁴, Zöe Powis, MD⁵, Kelly Minks, MS, CGC⁵, Aurélien Trimouille, MD^{6,7}, Eulalie Lasseaux, MD⁶, Didier Lacombe, MD, PhD^{6,7}, Chloé Angelini, MD^{6,7}, Vincent Michaud, MD^{6,7}, Julien Van-Gils, MD⁶, Nino Spataro, MSc⁷, Anna Ruiz, MD⁸, Elizabeth Gabau, MD⁹, Elliot Stolerman, MD¹⁰, Camerun Washington, CGC¹⁰, Raymond J. Louie, PhD¹⁰; Brendan C Lanpher, MD^{3,11}, Jennifer L. Kemppainen, MS, LCGC^{3,11}, A. Micheil Innes, MD¹², R. Frank Kooy, PhD¹³, Marije Meuwissen, MD, PhD¹³, Alice Goldenberg, MD¹⁴, François Lecoquierre, MD¹⁴, Gabriella Vera, MD, PhD¹⁴, Karin E M Diderich, MD, PhD¹⁵, Beth Rosen Sheidley, MS, CGC¹⁶, Christelle Moufawad El Achkar, MD¹⁶, Meredith Park, MS¹⁶, Fadi F. Hamdan, PhD¹⁷, Jacques L. Michaud, MD¹⁷, Ann J. Lewis, MD, PhD¹⁸, Christiane Zweier, MD, PhD^{19,20}, André Reis, MD²⁰, Matias Wagner, MD^{21,22}, Heike Weigand, MD²³, Hubert Journal, MD²⁴, Boris Keren, MD PhD^{25,26}, Sandrine Passemard, MD²⁷, Cyril Mignot, MD, PhD^{25,26}, Koen L.I. van Gassen, MD, PhD²⁸, Eva H. Brilstra, MD²⁸, Gina Itzikowitz, MD²⁹, Emily O’Heir, MS^{30,31}, Jake Allen, MS³², Kirsten A. Donald, MD, PhD^{33,34}, Bruce R. Korf, MD, PhD³⁵, Tammi Skelton, NP-C³⁵, Michelle L Thompson, MD^{36,37}, Nathaniel H. Robin, MD³⁶, Natasha Rudy, MS, CGC³⁶, William B. Dobyns, MD³⁶, Kimberly Foss, MS, LGC³⁶, Yuri A Zarate, MD³⁸, Katherine A. Bosanko, MS³⁸, Yves Alembik, MD³⁹, Benjamin Durand, MS³⁹, Frédéric Tran Mau-Them, MD, PhD¹, Emmanuelle Ranza, MD⁴⁰, Xavier Blanc, MD⁴⁰, Stylianos E. Antonarakis, MD, DSc⁴⁰, Kirsty McWalter, MS, CGC⁴¹, Erin Torti, MS, CGC⁴¹, Francisca Millan, MD, FACMG⁴¹, Amy Dameron, MS⁴¹, Mari J. Tokita, MD⁴¹, Michael T. Zimmermann, PhD^{42,43,44}, Nikita R. Dsouza, MS⁴², Eric W. Klee, PhD^{2,3,11}, Amélie Piton, PhD^{1,45,#}, Bénédicte Gerard, PharmD, PhD^{1,#}

¹Laboratoire de Diagnostic Génétique, Institut de génétique médicale d’Alsace (IGMA), Hôpitaux Universitaires de Strasbourg, Strasbourg, France

²Department of Health Sciences Research, Mayo Clinic, Rochester, MN, 55905, USA

³Center for Individualized Medicine, Mayo Clinic, Rochester, MN, 55905, United States

⁴Autism & Developmental Medicine Institute, Geisinger, Lewisburg, Pennsylvania, PA 17837, United States

⁵Department of Clinical Genomics, Ambry Genetics, Aliso Viejo, California, CA 92656, United States

⁶Service de Génétique Médicale, Centre de Référence Anomalies du Développement et Syndrome Malformatifs, CHU de Bordeaux, Bordeaux, France

⁷Maladies rares: Génétique et Métabolisme (MRGM), INSERM U1211, Université de Bordeaux, Bordeaux

⁸Genetics Laboratory, UDIAT-Centre Diagnòstic. Parc Taulí Hospital Universitari. Institut d’Investigació i Innovació Parc Taulí I3PT. Universitat Autònoma de Barcelona. Sabadell, Spain

⁹Paediatric Unit. Parc Taulí Hospital Universitari. Institut d’Investigació i Innovació Parc Taulí I3PT. Universitat Autònoma de Barcelona. Sabadell, Spain

¹⁰Greenwood Genetic Center, 106 Gregor Mendel Cir, Greenwood, SC 29646, USA

¹¹Department of Clinical Genomics, Mayo Clinic, Rochester, Minnesota, MN 55905, United States

¹²Department of Medical Genetics and Alberta Children’s Hospital Research Institute, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada

- ¹³Department of Medical Genetics, University and University Hospital Antwerp, Antwerp, Belgium
- ¹⁴Normandie Univ, UNIROUEN, Inserm U1245 and Rouen University Hospital, Department of Genetics and Reference Center for Developmental Disorders, F 76000, Normandy Center for Genomic and Personalized Medicine, Rouen, France
- ¹⁵Department of Clinical Genetics, Erasmus Medical Center, Rotterdam, The Netherlands
- ¹⁶Division of Epilepsy and Clinical Neurophysiology, Department of Neurology, Boston Children's Hospital, Boston, Massachusetts, MA 02115, United States
- ¹⁷Division of Medical Genetics, Department of Pediatrics, CHU Sainte-Justine and University of Montreal, Montreal, QC, Canada
- ¹⁸Pediatric Neurology, Kaiser Permanente Santa Clara Homestead, Santa Clara, United States
- ¹⁹Department of Human Genetics, Inselspital, Bern University Hospital, University of Bern, Bern, Switzerland
- ²⁰Institute of Human Genetics, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany
- ²¹Institute of Human Genetics, Technical University Munich, Munich, Germany
- ²²Institute of Neurogenomics, Helmholtz Zentrum München, Neuherberg, Germany
- ²³Department of Pediatric Neurology, Developmental Medicine and Social Pediatrics, Dr. von Hauner's Children's Hospital, University of Munich, Munich, Germany
- ²⁴Service de Génétique Médicale, Hôpital Chubert, Vannes, France
- ²⁵Département de Génétique et de Cytogénétique, Centre de Référence Déficience Intellectuelle de Causes Rares, GRC UPMC « Déficience Intellectuelle et Autisme », Hôpital Pitié-Salpêtrière, AP-HP, Paris, France
- ²⁶INSERM U 1127, CNRS UMR 7225, Sorbonne Universités, UPMC Univ Paris 06 UMR S 1127, Institut du Cerveau et de la Moelle épinière, ICM, Paris, France
- ²⁷Département de génétique, Hôpital Robert-Debré, AP-HP, Paris, France
- ²⁸Department of Genetics, Center for Molecular Medicine, University Medical Center Utrecht, Utrecht, The Netherlands
- ²⁹Department of Paediatrics and Child Health, Red Cross War Memorial Children's Hospital, University of Cape Town, SA
- ³⁰Center for Mendelian Genomics and Program in Medical and Population Genetics, Broad Institute of MIT and Harvard, Cambridge, MA, USA
- ³¹Division of Genetics and Genomics, Boston Children's Hospital, Boston, MA, USA
- ³²Stanley Center for Psychiatric Research, Broad Institute of MIT and Harvard, Cambridge, MA, USA
- ³³Department of Paediatrics and Child Health, Red Cross War Memorial Children's Hospital, University of Cape Town, SA
- ³⁴Neuroscience Institute, University of Cape Town, SA
- ³⁵Department of Genetics, University of Alabama at Birmingham, Birmingham, AL 35294, USA
- ³⁶Department of Pediatrics (Genetics) and Neurology, University of Washington, and Seattle Children's Research Institute, Seattle, Washington, USA
- ³⁷HudsonAlpha Institute for Biotechnology, Huntsville, Alabama, USA
- ³⁸Section of Genetics and Metabolism, University of Arkansas for Medical Sciences, Little Rock, USA
- ³⁹Service de Génétique Médicale, Institut de génétique médicale d'Alsace (IGMA), Hôpitaux Universitaires de Strasbourg, Strasbourg, France
- ⁴⁰Medigenome, Swiss Institute of Genomic Medicine, 1207 Geneva, Switzerland
- ⁴¹GeneDx, Gaithersburg, MD 20877, USA
- ⁴²Bioinformatics Research and Development Laboratory, Genomics Sciences and Precision Medicine Center, Medical College of Wisconsin, Milwaukee, WI 53226, USA
- ⁴³Clinical and Translational Sciences Institute, Medical College of Wisconsin, Milwaukee, WI 53226, USA
- ⁴⁴Department of Biochemistry, Medical College of Wisconsin, Milwaukee, WI 53226, USA
- ⁴⁵Institut de Genetique et de Biologie Moleculaire et Cellulaire, Illkirch 67400, France

CORRESPONDENCE:

Bénédicte Gerard, PharmD, PhD, Les laboratoires de diagnostic génétique, Hôpitaux Universitaires de Strasbourg, 1 place de l'Hôpital, 67000 Strasbourg, France. Email: benedicte.gerard@chru-strasbourg.fr

Amélie Piton, PhD, Laboratoire "Mécanismes génétiques des maladies neurodéveloppementales", IGBMC, 1, rue Laurent Fries, 67400, Illkirch, France. Email: piton@igbmc.fr

*These authors should be considered as equal junior contributors to this study

#These authors should be considered as equal senior contributors to this study

COMPETING INTEREST

Kirsty McWalter, Erin Torti, Francisca Millan, Amy Dameron, Mari J. Tokita are employees of GeneDx

ABSTRACT

High-impact pathogenic variants in more than 1,000 protein-coding genes cause Mendelian forms of neurodevelopmental disorders (NDD), including the newly reported *AGO2* gene. This study describes the molecular and clinical characterization of 28 probands with NDD harboring heterozygous *AGO1* coding variants. *De novo* status was always confirmed when parents were available (26/28). A total of 15 unique variants leading to amino acid changes or deletions were identified: 12 missense variants, two in-frame deletions of one codon, and one canonical splice variant leading to a deletion of two amino acid residues. Some variants were recurrently identified in several unrelated individuals: p.(Phe180del), p.(Leu190Pro), p.(Leu190Arg), p.(Gly199Ser), p.(Val254Ile) and p.(Glu376del). *AGO1* encodes the Argonaute 1 protein, which functions in gene-silencing pathways mediated by small non-coding RNAs. Three-dimensional protein structure predictions suggest that these variants might alter the flexibility of the AGO1 linkers domains, which likely would impair its function in mRNA processing. Affected individuals present with intellectual disability of varying severity, as well as speech and motor delay, autistic behavior and additional behavioral manifestations. Our study establishes that *de novo* coding variants in *AGO1* are involved in a novel monogenic form of NDD, highly similar to *AGO2* phenotype.

Keywords: *AGO1*, microRNA, intellectual disability, autism, argonaute

INTRODUCTION

Neurodevelopmental disorders (NDD), such as intellectual disability (ID) and autism spectrum disorder (ASD), have important genetic contributions characterized by extreme heterogeneity. More than a thousand genes have now been implicated in monogenic forms of NDD¹. Also, many more genes have been identified as candidates for ID/ASD, including genes showing enrichment of rare *de novo* variants in large-scale sequencing studies performed in affected individuals². In individuals affected by ID or ASD, few *de novo* missense variants have been reported in *AGO1* (or *EIF2C1*), a gene encoding a protein from the argonaute family, which participate in RNA silencing pathways, suggesting that *AGO1* could be a promising candidate gene for NDD³⁻⁶.

The argonaute protein family, identified originally in plants, includes AGO and PIWI proteins, and is involved in gene-silencing pathways guided by small non-coding RNAs (sncRNA, including short interfering RNAs, siRNAs; microRNAs, miRNAs; Piwi-interacting RNAs, piRNAs)⁷. PIWI proteins are involved in transposon repression in germinal cells whereas AGO proteins are involved in translation repression and degradation of targeted mRNA⁸. In addition to their role in mRNA post-transcriptional regulation in the cytoplasm, AGO proteins have also been shown to have nuclear activities, playing a role in the regulation of transcription, chromatin remodeling, alternative splicing regulation, and even in DNA double-strand break repair⁸⁻¹².

Large deletions of 1.1 Mb to 3.1 Mb at the 1p34.3 loci including *AGO1* together with *AGO3* (and sometimes *AGO4* among other genes) were previously reported in children with NDD^{13,14}. These five individuals presented with psychomotor developmental delay as well as additional non-specific features (feeding difficulty, language impairment, facial dysmorphism). In addition, *AGO2* was very recently implicated in NDD¹⁵: 21 individuals with heterozygous *de novo* mutations in *AGO2* variant were reported, including 11 missense variants, one in-frame deletion and one 235 kb deletion involving the first three exons of *AGO2*. Functional studies revealed that those variants hampered correct sncRNA-mediated silencing, having a loss of function effect.

We here report 28 individuals from 26 families affected by NDD carrying heterozygous amino acid substitutions or deletions predicted-damaging in *AGO1*, thus, causally-linking coding variants in this gene to ID. These variants distributed along the different domains of the protein, are predicted to affect the flexibility of AGO1 structure, which might impair its role in sncRNA-

induced gene regulation. Several *AGO1* variants reported in this study affect homologous residues to those recently described in *AGO2* and faithfully recapitulate the functional prediction effects. The high similarity in clinical presentation reinforces the description of *AGO1/AGO2* associated phenotype.

MATERIALS AND METHODS

Cohort

All affected individuals were recruited independently through a worldwide collaborative network of clinical and molecular geneticists connected notably via GeneMatcher¹⁶. Affected individuals were referred by clinical genetics services from across Belgium (family 8), Canada (families 7, 12), France (families 3, 9, 16, 17, 20 and 24), Germany (families 14 and 15), Switzerland (family 26), USA (families 1, 2, 5, 6, 11, 13, 19, 21, 22, 23 and 25), Spain (family 4) and The Netherlands (families 10 and 18). Blood samples were obtained following the provision of informed consent by the individuals or their legal representatives. Ethical approval was obtained from the local ethics committees.

Identification of AGO1 variants

AGO1 variants were identified by laboratories using various high throughput sequencing strategies (targeted genes panel, exome sequencing or whole genome sequencing) (see **Table 1**). Reads alignment was performed against the human genome reference build GRCh37/hg19; Filtering strategies and description of bioinformatics tools are presented in supplementary data (**Supplementary text**). *AGO1* variants were annotated following the NM_012199.4. Variants were confirmed by independent Sanger sequencing and concordance of the trio (father, mother, child) was confirmed for all *de novo* variants, either by verifying rare SNP inheritance pattern in the case of NGS trios or by microsatellites segregation study, except for families 23 and 25, as no parental samples were available. Predictions of missense variant effects were performed using *in silico* tools such as Combined Annotation Dependent Depletion (CADD¹⁷), SIFT¹⁸ and Polyphen-2¹⁹. Effects on splicing were analyzed using SpliceAI²⁰, NNsplice²¹ and MaxEnt²² and confirmed, if needed, by cDNA analysis from blood mRNA.

Collection of clinical information

Physicians with experience in clinical dysmorphology examined *a posteriori* all 28 individuals of this cohort, including the individual carrying the previously described variant in Hamdan *et al*⁵ (F12). Photographs were collected from research participants after informed consent had been obtained. Clinical data for additional individuals with *de novo* variants reported in previous publications^{3-6,23,24} were retrieved from published data or supplemental data. Face2Gene Research application (FDNA Inc., Boston, MA) using DeepGestalt technology (algorithm 19.1.9)²⁵ was used to evaluate the facial characteristics of the individuals with *AGO1* variants. Fourteen frontal facial photos were obtained from unrelated affected individuals (proband family: age at photo: F3: 7 years, F4: 8 years, F5: ~3-6 years, F6: 12 years, F9 twin 2: 3 years, F10: 8 years, F11: 11 years, F13: 7 years, F14: 5.7 years, F19: 2 years, F20: 3 years, F21: 3.5 years, F23: 9 years, F24 twin1: 10 years) and compared to 14 control images matched for age, sex, and ethnicity provided by Face2Gene. To estimate the power of DeepGestalt in distinguishing affected individuals from controls, a cross validation scheme was used, including a series of binary comparisons between all groups. For these binary comparisons, the data were split randomly multiple times into training sets and test sets. Each set contained half of the samples from the group, and this random process was repeated 10 times. The results of the comparisons are reported using the receiver operating characteristic (ROC) curve and area under the curve (AUC).

Building 3D Model for AGO1

The protein structure of human AGO1 has been experimentally determined (PDB 4kxt²⁶). We used the available experimental structure with homology-based methods²⁷ to fill in short sections of the protein that were not resolved in the experiment. Structures were solved with simple (poly A) guide RNAs that were also partially resolved. We added data from the *Rhodobacter sphaeroides* Argonaute experimental structure (PDB 5awh²⁸) and used Discovery Studio (BIOVIA, Dassault Systèmes BIOVIA, Discovery Studio Modeling Environment, Release 4.5, San Diego: Dassault Systèmes; 2015) to complete coordinates of a 21-base (A₁₉U₂) guide RNA. We used BioR²⁹ to compile protein annotations from multiple sources including dbNSPF³⁰. We used FoldX³¹ v4.0 for mutagenesis and structure-based calculations of $\Delta\Delta G_{\text{fold}}$.

We visualized protein structures using PyMOL (Molecular Graphics System, Version 1.9 Schrödinger, LLC).

Molecular Dynamics Simulations

Generalized Born implicit solvent molecular dynamics (isMD) simulations were carried out using NAMD³² and the CHARMM36 force field, using a similar procedure to our previous work. Briefly, we utilized an interaction cutoff of 12Å with strength tapering (switching) beginning at 10Å, a simulation time step of 1fs, conformations recorded every 2ps. Each initial conformation was used to generate 6 replicates, and each was energy minimized for 5,000 steps, followed by heating to 300K over 300ps via a Langevin thermostat. A further 13ns of simulation trajectory was generated and the final 10ns (60ns per variant) were analyzed.

Analysis of Protein Structures and Simulations

All trajectories were first aligned to the initial wild type conformation using C^α atoms. Root Mean-Square Deviation (RMSD) values were reported for each after aligning to the initial WT conformation. We calculated per-domain RMSDs by first re-aligning all trajectories using each domain individually, then measuring the RMSD within that domain, alone. This was done independently for each domain, providing a measure of distortion within each domain and across MD simulation time. Root Mean-Square Fluctuation (RMSF) values were calculated at the residue level for the whole protein across trajectories aligned to the initial WT conformation. Variances were computed using median absolute difference (MAD). We computed Z-scores of data by MAD-scaling the difference between each observation and the median. We used C^α cartesian space Principal Component (PC) analyses across all simulations to define the essential dynamics of AGO1. Data from each variant was projected onto the PCs to compare how essential motion is activated or suppressed by each genomic variant. Individual PCs were visualized using porcupine plots where a cone is placed to represent the direction and relative magnitude of each residue's motion. We calculated free-energy landscapes (FELs) of MD trajectories using the approach of Karamzadeh, *et al.*³³. We show topologic lines from the FEL where each line indicates a specific conformational sampling probability and matched by a corresponding color gradient where each color also indicates a specific conformational sampling probability. We calculated internal distances for an alignment-free measure of

conformational changes³⁴. We used the ensemble-averaged median pairwise distance between residues, calculated on the last half of each trajectory, as a summary of the change in internal distances. To simplify visualization of these median changes, we averaged information across groups of three consecutive amino acids. To conservatively estimate statistical significance, we used a permutation procedure where data were sub-sampled to 100 points and compared using a t-test. This was repeated 1000 times and the median p-value across repeats reported. The analysis was carried out using custom scripts, leveraging VMD and the Bio3D R package. Protein structure visualization was performed in PyMol and VMD.

RESULTS

Identification of de novo variants in AGO1 in individuals with intellectual disability

A total of 15 different heterozygous variants in *AGO1*, including 12 amino acid substitutions, two single amino acid deletions, and one splice variant leading to a two-residues deletion were identified in 26 unrelated families (**Table 1, Figure 1**). For all individuals for whom parental DNA was available (all families except 23 and 25), variants were found to occur *de novo*. No additional genetic variant potentially explaining the clinical manifestations was identified in all these individuals (**Supplementary text**). Two of the amino acid changes we identified had already been reported in individuals with NDD: p.(Leu190Pro) (Rauch *et al.*³) and p.(Gly199Ser) (Hamdan *et al.*⁵ and Sakaguchi *et al.*⁶). Thirteen variants were novel to this study. The proband from family F15 has a *de novo* c.650-2A>G variant predicted to impair the use of the canonical acceptor splice site (MaxEnt: -100.0%, NnSplice: -100.0%) and to activate a cryptic acceptor site (Splice AI prediction). Activation of the cryptic acceptor site was subsequently confirmed by mRNA Sanger sequencing (**Supplementary Figure S1**), leading to the in-frame deletion of two amino acids (r.650_655del, p.Val217_Ser218). Six variants were recurrent: p.(Gly199Ser) (six individuals), p.(Phe180del) (five individuals), p.(Leu190Arg) (two individuals), p.(Val254Ile)(two individuals) and p.(Glu376del) (two individuals). The remaining variants were unique: p.(Pro189Leu), p.(Arg253His), p.(Val217_Ser218del), p.(Arg253His), p.(Pro324Leu), p.(Gln358Arg), p.(Tyr418Phe), p.(Thr781Met) and p.(Ile797Phe). Including the three previously reported missense variants: p.(Glu195Lys)²³, p.(Asp216Val)²⁴ and p.(Thr355Ile)⁴, a total of 18 variants were identified in individuals with NDD (**Table 1, Figure 1**). All identified variants were absent from the general population included in the gnomAD database and all but one affect amino acid residues that are highly

conserved until *C. elegans* (**Supplementary Figure S1**). *In silico* predictions using SIFT and Polyphen2 predict all these missense variants to be deleterious except p.(Arg253His) and p.(Val254Ile). A depletion of missense changes is observed in general populations, with three times fewer missense variants observed than expected in gnomAD cohorts (ratio=0.31, Z-score= 5.68). The CADD scores of missense variants identified in individuals with NDD (mean=27.8) are significantly higher than those of missense variants reported in gnomAD (mean: 24.3, t-test p-value=0.0001).

Clinical manifestations observed in individuals with AGO1 variants

We collected the clinical characteristics of the 28 affected individuals enrolled in this study: first clinical description for 27 individuals and updated clinical information for one previously reported⁵. We also retrieved the clinical information from previous publications reporting individuals with variants in *AGO1*^{3-6,23,24} and thus obtained clinical data for a total of 33 individuals (17 males, 16 females) from 31 unrelated families. Head circumference and height and weight measurements were in the normal ranges at birth and postnatally, except for three individuals (including two twins) who showed neonatal and postnatal microcephaly (<-2 SD). All affected individuals showed borderline to severe ID, (**Table 2, Supplementary text**) and all showed a severe language delay (30/30; 100 %). Most individuals were able to construct a sentence but with limited spontaneous communication. Three did not acquire language at all and one presented with language regression. A motor developmental delay was also observed in most of the affected individuals: 93 % of individuals had delayed walking (28/30; mean age at onset of independent walking ~25 months) and seven individuals had persistent hypotonia (13/18; 72 %). Almost half of the individuals had documented seizures or history of seizures (13/28; 46 %), seizures were fever-induced in two families (F7, F18). Twin girls (F9) have different types of seizure: one has photosensitive, drug-resistant seizure and the other has controlled seizure. Interestingly, the three individuals with the most severe epileptic phenotype (as indicated by the history of status epilepticus) share the same variant p.(Gly199Ser) (the fourth individual with this variant has no history of epilepsy at the age of 5 years). Several behavioral features were observed in patients with *AGO1* variants: most of the individuals presented with autistic features (24/30; 80 %), and 11 individuals showed self-harm behavior and/or hetero-aggressiveness (11/14, 78.5%). Attention deficit/hyperactivity was also noticed (15/22, 68 %) and anxiety was reported in seven individuals (7/8, 87.5 %). Moreover, a high

percentage of individuals had various sleeping disturbances (17/22, 77 %) including difficulties falling asleep, early awakening, or hypersomnia. Analysis using DeepGestalt technology by Face2Gene did not show significant differences in facial features from matched controls (mean AUC = 0.55 and AUC STD = 0.14, p-value = 0.392) (**Figure 2**). However, if they do not present a recurrent recognizable gestalt, individuals with *AGO1* variants share some subtle common traits such as a thin upper lip, a tall forehead, elongated almond eyes or a small nose. Individuals with feeding difficulties (10/23, 43 %) showed low postnatal weight: a gastrostomy was necessary for six individuals, three individuals had gastro-esophageal reflux and the twins of family 24 received tube feeding during the first three weeks of life in the context of prematurity. Two individuals had recent swallowing problems (F3 and F25). Finally, hypothyroidism was reported in two out of five individuals with the p.(Phe180del) variant. Variable brain MRI anomalies were noticed in 11/24 individuals (46 %), including corpus callosum anomalies (agenesis or hypoplasia), cerebral atrophy, cortical dysplasia or colpocephaly but also non-specific anomalies (increase in extra-axial fluid, T2 and FLAIR signal anomalies, arachnoid cyst).

Variants spatially cluster in AGO1

The AGO1 protein (NP_036331.1) contains five identified functional domains: two linker domains L1 (AA 173-225) and L2 (AA 372-418), one MID domain (AA 427-508), one PAZ domain (AA 226-368) and one PIWI domain (AA 515-816). The PAZ domain contains the RNA binding module that recognizes the 3' end of both siRNA and miRNA. The PIWI domain contains the cleavage site that is inactive in AGO1 and is involved in protein/protein interactions, notably with DICER and GW182. The nonsynonymous variants cluster in 17 amino acid positions of the protein: eight are in the L1 binding domain (p.Phe180, p.Pro189, p.Leu190, p.Glu195, p.Gly199, p.Asp216, p.Val217_Ser218), four in the PAZ domain (p.Arg253, p.Val254, p.Pro324, p.Thr355, p.Gln358), two variants in the L2 domain (p.Glu376, p.Tyr418) and three in the PIWI domain (p.His751, p.Thr781, p.Ile797) (**Table 2**).

The majority of 10 first collected variants affect amino acids (Phe180, Pro189, Leu190, Glu195, Gly199, Asp216, Arg253, Val254 and His751) clustered in 3D along the sides of L1 and PAZ domains facing one another in 3D, except for His751 (**Figure 3A**). Indeed, within the PAZ domain, Arg253 is within an alpha-helix and makes specific contact with the last residue in the helix, Asp250, and the backbone of Arg284. Thus, polar contacts involving Arg253 are likely

important for the organization and stability of the PAZ domain. The variant p.(Arg253His) may modify these interactions. The nearby variant p.(Val254Ile) is conservative, but valine has the lowest helical propensity of the small hydrophobic amino acids³⁵. Thus, this variant could over-stabilize the helix and limit motion of the PAZ domain. Three critical amino acid positions are within the L1 domain and in proximity to the PAZ domain: Pro189, Leu190, and Glu195 (**Figure 3B**). Pro189 is at the base of a long loop and the backbone geometry of proline may be necessary for limiting the loop's mobility; p.(Pro189Leu) may thus alter loop dynamics. Leu190 packs within a hydrophobic interface made up of amino acids from L1 and L2; p.(Leu190Pro) or p.(Leu190Arg) may distort the interface between the linker domains. Glu195 is close in space to Trp197 and Lys224; p.(Glu195Lys) may repel these nearby amino acids and destabilize the L1 domain. Farther down the L1 domain, p.(Gly199Ser) is a recurrent variant (observed in four unrelated individuals and published once in another individual): the side chain of the introduced serine could clash with the side chain of Ser218 and introduce an unfavorable polar contact, likely destabilizing the L1 domain.

Phe180 makes contacts across all three strands of the L1 beta sheet and its deletion would significantly alter the hydrophobic packing of the region. Asp216 makes a salt bridge with Arg712 in the PIWI domain (**Figure 3C**). Arg712 also makes specific interactions with the guide RNA. Thus, the variant p.(Asp216Val) is likely to destabilize the interface between L1 and PIWI, and thereby indirectly alter interactions with the guide RNA. Finally, p.(His751Leu) is the only observed variant within the PIWI domain. His751 makes hydrogen bonds with the guide RNA and p.(His751Leu) is therefore likely to alter the stability of the interaction between AGO1 and the guide RNA.

Variants are predicted to affect AGO1 dynamic conformational change

We used our molecular dynamic (MD) simulations to assess global and local changes to AGO1 in association with each genomic variant. We first assessed global changes to AGO1 conformation across our MD simulations using RMSD, RMSF and PC analysis (**Figure 4**). We found that genomic variants have the largest change in the patterns of mobility for the N-term and PAZ domains, and more modest mobility changes to the MID domain (**Figure 4A**). Because RMSF depends on how the trajectories are aligned, and we aligned based on the PIWI domain, we observed relatively small overall fluctuations of the PIWI domain. Therefore,

patterns in RMSF indicated that the largest features in our simulations were domain-domain motions, and that these were modified by genomic variants.

Next, to assess domain-domain distance, we assessed linear distances from specific residues in the PIWI, MID, and linker regions. Certain PAZ and linker variants such as p.(Val254Ile), p.(Gly199Ser), p.(Arg253His), and p.(Pro189Leu) were associated with significantly shorter distances between the PIWI and linker (**Figure 4B**). The same variants p.(Val254Ile) and p.(Gly199Ser) and additional ones p.(Tyr418Phe) and p.(His751Leu) were associated with significantly shorter distances between the MID and linker (**Figure 4C**). Because the relative orientation of these domains is likely important for AGO1 function, we also assessed the angles between them. Genomic variants in multiple domains altered the orientation of PAZ with respect to linker (**Figure 4D**), and the MID-PIWI-PAZ orientation (**Figure 4E**). Thus, we identified additional and more specific changes in domain-domain orientations associated with genomic variants.

Because many conformational changes occur concurrently in our MD simulations, we next assessed all dynamics data together using PC analysis (**Figure 4F**). We visualized the predominant dynamics of AGO1 using a free-energy landscape. We found that the WT protein sampled one side of the landscape, while variants had significant sampling of a different side. PC states are characterized by the relative orientations of the PAZ, Nterm, and MID domains (see **Supplemental Animation 1** for PC visualizations). Genomic variants had two effects on these predominant dynamics. First, they shift the average conformation to one that favors a wider angle of the PAZ-PIWI-MID domains. Second, the transition between conformations was blurred with a significant amount of time that the protein deviates from WT-like conformations. This indicates that the genomic variants may change the nature of AGO1, leading to dysregulated dynamics. We explored the possibility for changes in internal distances as a summary of dysregulated dynamics (**Supplementary Figure S3**) and demonstrated that genomic variants were associated with lower coordination between domains. We anticipate that changes to domain coordination could be an additional and informative criterion for assessing multi-domain enzymes.

To further assess the intra-domain change associated with each amino acid variant, we calculated the per-domain RMSD (**Supplementary Figure S4**). We found that amino acid changes throughout the structure could be associated with alterations of the same or distant domains. For example, p.(Leu190Pro) is within the linker domain and the conservative variant

p.(Val254Ile) is within the PAZ domain but both lead to alterations in the internal organization of the linker and MID domains.

DISCUSSION

This study presents the clinical and molecular characterization of a cohort of 28 individuals from 26 unrelated families with 15 unique heterozygous coding variants (13 missense and two in-frame deletion variants) in *AGO1*. This study establishes that *de novo* coding variants in *AGO1* are responsible for a form of NDD characterized by psychomotor delay, behavioral features and language impairment. Including the three previously reported variants, a total of 18 variants have been described in *AGO1* in a total of 33 individuals with NDD. Six variants are recurrent: p.(Gly199Ser), p.(Phe180del), p.(Leu190Pro), p.(Leu190Arg), p.(Val254Ile), p.(Glu376del), while the others are reported in only one family to date. All the variants lead to substitutions or deletions of one or two amino acids, but no truncating variant (nonsense, frameshift, or splicing variants leading to truncation) was observed in this cohort. Interestingly, four of the amino acids residues mutated in *AGO1*, Phe180, Leu190, Gly199 and Thr355, were also found mutated at the equivalent residues in *AGO2* (p.Phe182del, p.Leu192Pro, p.Gly201Cys or Val, p.Thr357Met) in ID patients (**Supplementary Figure S1**). Others *AGO2* missense variants were reported in L1 domain (p.His203Gln), in helix-7 of the L2 domain (p.Met364Thr, p.Ala367Pro) involved in the proper positioning of the RNA guide, in the MID domain (p.Gly573Ser), and in PIWI domain (p.Gly733Arg, p.Cys751Tyr and p.Ser760Arg). The last three variants are located in a loop which binds in the minor groove of guide/target duplex¹⁵.

ID and language delay were reported in all affected while most of them displayed motor delay, seizures, autistic features, and behavior problems. No recurrent structural brain nor other organs malformation was noticeable and if individuals share some common facial traits, no typical specific gestalt was observed. Therefore, it appears that the NDD form caused by coding variants in *AGO1* is not a clearly recognizable entity. Moreover, the individuals present differences in the severity of the clinical manifestations, which do not appear to correlate with the nature or the location of the variant. The severity of ID, as well as growth parameters, was for instance highly variable among the three individuals with the p.(Phe180del) variant. We could therefore suggest that additional genetic or environmental factors might play a role in the phenotypic expressivity, modulating the effect of the *AGO1* variant. On the contrary, some

specific traits seem to correlate with specific variants, such as hypothyroidism specifically reported in two individuals with the p.(Phe180del) variant, or status epilepticus in three individuals with p.(Gly199Ser) variant. A larger number of individuals would be necessary to confirm these observations. Individuals with *AGO1* or *AGO2* variants showed common clinical findings: ID with autistic features and aggressiveness, impaired speech development, motor delay, MRI anomalies and frequent gastrointestinal disorder or reflux. Minor additional clinical features are also reported in both cohorts: skeletal (clinodactyly or bradymetacarpus, scoliosis), vision problem (strabismus, myopia/hyperopia, visual impairment), heart, dental or breathing anomalies, and also anxiety and sleeping disturbance. Frequency of such minor anomalies should be confirmed by future studies. To note, we observed three monozygotic couples in *AGO1/AGO2* cohorts. This apparent excess of monozygotic twins (6/54 patients) have to be confirmed with future studies.

The 3D conformational model of AGO revealed a very flexible protein with two mobile domains, the L1 and L2 linker domains. The global structure contains four globular domains (PAZ, MID, PIWI and N terminal), and a deep cleft bordered by L1 and L2 linkers. L1 and PAZ domains seemed to be more flexible than the L2, PIWI and MID domain, as shown in **Figure 3**. The flexibility of the protein seems to be important to permit transitions between the different phases of the RISC process: sncRNA loading and processing, mRNA target clip, helper protein recruitment, and finally RISC complex release. Opening and closure of the PAZ/L1 jaw seems necessary for the proper function of AGO protein. Among the 18 *AGO1* variants reported to date the majority clustered within the L1, PAZ and PIWI domains. We have analyzed the potential effects of variants located in the L1 and L2 linker domains, as well as some of the variants located in the PAZ and PIWI domains, using structural biology as an interpretive framework. We considered the role of native residues within *AGO1* and how they would be altered by the variants. Our assessment of the *AGO1* variants identified in individuals with NDD indicated that they were deeply buried in the cleft region and close to the RNA guide molecule, but did not affect the residues involved in RNA guide anchorage; in addition, dynamical simulations showed that these variants narrowed the angulation between the PAZ/L1 domains. These observations suggested these variants may hamper *AGO1* flexibility during the different phases of the mRNA processing. Recently, the *AGO1x* protein isoform was described, product of the translational stop read-through of *AGO1* transcripts induced by the expression of the Let7a miRNA. Ectopic expression of *AGO1x* acts as a negative competitor of the miRNA pathway³⁶. A hypothesis would be that mutated *AGO1* proteins may act as the *AGO1x* isoform

through competitive inhibition of wild type AGO proteins. Functional studies will be necessary to investigate this hypothesis.

The absence of truncating *AGO1* variants in our cohort is surprising and is not in favor of a loss-of-function (LoF) as the unique molecular mechanism involved in *AGO1*-related ID. However, the observation of five children with large heterozygous deletions including *AGO1* together with *AGO3* (and sometimes *AGO4* among other genes), and a deletion encompassing the *AGO2* first 3 exons challenges this observation^{13–15}. *AGO2* variant functional analysis revealed a complex cellular deregulation: a decrease in mRNA silencing was observed as expected for a LoF mechanism but with a variable impact, depending on the mRNA target and the tested *AGO2* variant. Reduced target release and reduced phosphorylation of the serine cluster at residues 824 to 834 in *AGO2* was observed for most of the tested variants due to probable protein dynamics perturbation, likely leading to *AGO2* deregulation and reduced functions¹⁵.

A better understanding of how *AGO1* functions in the brain will be essential to understanding how pathogenic variants in this gene cause cognitive impairment, developmental delays and behavioral manifestations including autistic traits. In the developing mouse brain, although it is ubiquitously expressed, *Ago1* was shown to be upregulated in neurogenic progenitors and mature neurons³⁷. This expression pattern correlates with initial findings showing a role for human *AGO1* in promoting differentiation after neuronal induction in a cellular model of neuroblastoma SH-SY-5Y cells³⁷. Mouse models carrying homozygous inactivation of *Ago1* showed postnatal lethality as reported in IMPC (International Mouse Phenotyping Consortium, mousephenotype.org): only 9 % of homozygotes *Ago1*^{tm1a/tm1a} mice are obtained from heterozygous mating suggesting a prenatal lethality. The surviving homozygotes (both males and females) showed decreased body weight and increased anxiety (abnormal behavior in the open field). Pathogenic variants in several other genes encoding protein partners of *AGO1/AGO2* involved in the regulation of mRNA decay and translation, such as *CNOT1*³⁸, *CNOT2*³⁹, *CNOT3*⁴⁰ or *DDX6*⁴¹, have been reported to cause NDDs⁴². Apart from their role in post-transcriptional regulation, human *AGO1* and *AGO2* proteins are also involved in the regulation of transcription and splicing⁹. Therefore, alterations of these nuclear functions could also be involved in the pathophysiology of the neurodevelopmental disorders caused by coding variants in these genes.

In conclusion, this collaborative study, reporting molecular and clinical data from 26 families with heterozygous *de novo* coding variants in *AGO1*, confirms the involvement of rare coding variants in this gene in NDD. Future studies investigating transcriptional and posttranscriptional regulation defects or other dysfunction stemming from *AGO1* pathogenic variants will be important for elucidating the precise mechanisms of disease that may inform potential therapeutic strategies.

ACKNOWLEDGEMENTS

We thank all the families for their participation in the study. We also thank the CREGEMES for its financial support. The authors want also to thank all the people from the Strasbourg Hospital molecular diagnostic lab, especially the team working on diagnosis of Intellectual Disability (Claire Feger, Elsa Nourisson, Céline Cuny, Sylvie Friedmann and Carmen Fruchart) and all the bioinformatics team. The authors want also to thank all the people from the Cambridge Broad Institute of MIT and Harvard, especially the team working on clinic and diagnostic of patients (Victoria de Menil, Elise Robinson, Patricia Kipkemoi and Anne O'Donnell Luria). We also thank the Centre National de Génotypage (Jean-François Deleuze, Robert Olaso, Anne Boland, and the technicians and bioinformaticians) for their participation in library preparation and DNA sequencing. We also thank the Center for Individualized Medicine at Mayo Clinic for their support. FL, GV and AG received funding from European Union and Région Normandie in the context of Recherche Innovation Normandie (RIN 2018). Europe gets involved in Normandie with the European Regional Development Fund (ERDF).

WEB RESOURCES

CADD, <https://cadd.gs.washington.edu/snv>

GeneMatcher, <https://genematcher.org/>

GnomAD browser, <https://gnomad.broadinstitute.org/>

OMIM, <https://www.omim.org/>

SySID, <https://sysid.cmbi.umcn.nl/>

FIGURE LEGENDS

Figure 1: Schematic representation of AGO1 protein with its functional domains showing the locations of the missense variants identified in individuals with ID

AGO1 protein (NP_036331.1) with its functional domains: N-terminal domain (34-164), Linker 1 (L1) domain (173-225), PAZ domain (226-368), Linker 2 (372-418), Mid domain (427-508), and PIWI domain (515-816). PAZ and PIWI domains have motifs of interaction with

RNA guide and PIWI domain has motif of RNA blocked access to the active site. The arrows show *de novo* variants identified in individuals from this cohort or previously reported (P: previously reported in literature); T: found in monozygotic twins, and in underlined those that are recurrent.

Figure 2: Facial characteristics of individuals with *AGO1* variants.

(A) Front faces, (B) profile faces (C) Face2Gene Facial analysis using Face2Gene Research application (FDNA Inc. Boston, MA) of unrelated individuals with *AGO1*-associated disorder (n = 14) compared to controls matched (n= 14) for sex, age, and ethnicity.

Figure 3: *AGO1* structure and annotation of novel variants.

(A) Protein structure of *AGO1* colored by protein domain and with the sites of variants indicated by spheres. The linker domains, designated L1 and L2, are separated in sequence by the PAZ domain, but intertwine in 3D, forming common interfaces between the N-terminal, PAZ, and PIWI domains (B) Many of the variants of interest are within the first linker domain. This domain forms a narrow beta sheet with three strands. Variants occur in the middle of these strands at the closest point between domain L1 and the guide RNA backbone, and towards the end of L1 and near the PAZ domain interface (C) D216 is within L1 and makes specific contacts with R712 in the PIWI domain. R712 also interacts directly with the guide RNA.

Figure 4: Simulations reveal changes in domain orientation associated with *de novo* missense *AGO1* variants

We used MD simulations to assess how the native structure of *AGO1* would respond to the introduction of a subset of identified genomic variants. (A) Variability of each amino acid was quantified using RMSF after aligning each trajectory to the initial WT conformation of the PIWI domain and averaging across replicates. Domains are colored as in Figure 1 and each variant colored according to the domain it is within. (B-E) We monitored selected distances and angles as a simple way to assess conformational changes between the (B) linker and PIWI domains, (C) linker and MID domains, (D) the orientation of the PAZ domain, and (E) the openness of the RNA binding region. (F) We show the free energy landscape across molecular

dynamics (MD) simulations as a color gradient from high-energy to low energy. Above, we show one-dimensional PC samplings as a combined violin and boxplot. The left- and right-hand panels summarize all data from the WT and from our novel variants, respectively. Selected variant's PC1 sampling is shown above the panel, and all variant's (that underwent MD) PC2 sampling to the side of the panel. Four regions of low energy are indicated by letters A-D. Representative images of AGO1 structure from the simulations taken from these points of low energy are shown below and labeled by the corresponding letters with the WT shown for comparison. To summarize one aspect of the difference between these four regions, we show the angle between the centers of mass of the MID, PIWI, and PAZ domains.

Table 1: List of missense changes identified in *AGO1* in individuals with intellectual disability

Version of the programs used: SIFT version v6.2.0; Polyphen2 (PPH2) HumVar version v2.2.2r398; CADD version GRCh37-v1.4; †Change in folding free energy upon each mutation, using a 3D structure-based algorithm. See Methods for details. ‡3D molecular context of the amino acid position. See Results for details. αMD simulations were summarized using PC analysis; we summarized the change in PC1 and PC2 using the number of standard deviations away from the average WT conformation that each mutation sampled on average.

Table 2: Clinical manifestations observed in individuals with variants in *AGO1* (compared to individuals with variants in *AGO2* reported by Lessel et al. 2020)

ID: intellectual disability; DD: developmental delay; M: male; F: female; NA: not reported

SUPPLEMENTAL DATA

REFERENCES

1. Vissers LELM, Gilissen C, Veltman JA. Genetic studies in intellectual disability and related disorders. *Nature Reviews Genetics*. 2016;17(1):9-18. doi:10.1038/nrg3999
2. Lelieveld SH, Wiel L, Venselaar H, et al. Spatial Clustering of de Novo Missense Mutations Identifies Candidate Neurodevelopmental Disorder-Associated Genes. *Am J Hum Genet*. 2017;101(3):478-484. doi:10.1016/j.ajhg.2017.08.004
3. Rauch A, Wieczorek D, Graf E, et al. Range of genetic mutations associated with severe non-syndromic sporadic intellectual disability: an exome sequencing study. *The Lancet*. 2012;380(9854):1674-1682. doi:10.1016/S0140-6736(12)61480-9
4. Sanders SJ, Murtha MT, Gupta AR, et al. De novo mutations revealed by whole exome sequencing are strongly associated with autism. *Nature*. 2012;485(7397):237-241. doi:10.1038/nature10945
5. Hamdan FF, Srouf M, Capo-Chichi J-M, et al. De Novo Mutations in Moderate or Severe Intellectual Disability. Cooper GM, ed. *PLoS Genetics*. 2014;10(10):e1004772. doi:10.1371/journal.pgen.1004772
6. Sakaguchi A, Yamashita Y, Ishii T, et al. Further evidence of a causal association between AGO1, a critical regulator of microRNA formation, and intellectual disability/autism spectrum disorder. *European Journal of Medical Genetics*. 2019;62(6):103537. doi:10.1016/j.ejmg.2018.09.004
7. Sasaki T, Shiohama A, Minoshima S, Shimizu N. Identification of eight members of the Argonaute family in the human genome☆ *Genomics*. 2003;82(3):323-330. doi:10.1016/S0888-7543(03)00129-0
8. Meister G. Argonaute proteins: functional insights and emerging roles. *Nat Rev Genet*. 2013;14(7):447-459. doi:10.1038/nrg3462
9. Allo M, Agirre E, Bessonov S, et al. Argonaute-1 binds transcriptional enhancers and controls constitutive and alternative splicing in human cells. *Proceedings of the National Academy of Sciences*. 2014;111(44):15622-15629. doi:10.1073/pnas.1416858111
10. Shuaib M, Parsi KM, Thimma M, et al. Nuclear AGO1 Regulates Gene Expression by Affecting Chromatin Architecture in Human Cells. *Cell Systems*. 2019;9(5):446-458.e6. doi:10.1016/j.cels.2019.09.005
11. Ameyar-Zazoua M, Rachez C, Souidi M, et al. Argonaute proteins couple chromatin silencing to alternative splicing. *Nature Structural & Molecular Biology*. 2012;19(10):998-1004. doi:10.1038/nsmb.2373
12. Wei W, Ba Z, Gao M, et al. A role for small RNAs in DNA double-strand break repair. *Cell*. 2012;149(1):101-112. doi:10.1016/j.cell.2012.03.002
13. Tokita MJ, Chow PM, Mirzaa G, et al. Five children with deletions of 1p34.3 encompassing AGO1 and AGO3. *European Journal of Human Genetics*. 2015;23(6):761-765. doi:10.1038/ejhg.2014.202
14. Jacher JE, Innis JW. Interstitial microdeletion of the 1p34.3p34.2 region. *Mol Genet Genomic Med*. 2018;6(4):673-677. doi:10.1002/mgg3.409

15. Lessel D, Zeitler DM, Reijnders MRF, et al. Germline AGO2 mutations impair RNA interference and human neurological development. *Nat Commun.* 2020;11(1):5797. doi:10.1038/s41467-020-19572-5
16. Sobreira N, Schiettecatte F, Valle D, Hamosh A. GeneMatcher: A Matching Tool for Connecting Investigators with an Interest in the Same Gene. *Human Mutation.* 2015;36(10):928-930. doi:10.1002/humu.22844
17. Kircher M, Witten DM, Jain P, O’Roak BJ, Cooper GM, Shendure J. A general framework for estimating the relative pathogenicity of human genetic variants. *Nature Genetics.* 2014;46(3):310-315. doi:10.1038/ng.2892
18. Ng PC, Henikoff S. SIFT: Predicting amino acid changes that affect protein function. *Nucleic Acids Res.* 2003;31(13):3812-3814.
19. Adzhubei I, Jordan DM, Sunyaev SR. Predicting Functional Effect of Human Missense Mutations Using PolyPhen-2. *Curr Protoc Hum Genet.* 2013;07:Unit7.20. doi:10.1002/0471142905.hg0720s76
20. Jaganathan K, Kyriazopoulou Panagiotopoulou S, McRae JF, et al. Predicting Splicing from Primary Sequence with Deep Learning. *Cell.* 2019;176(3):535-548.e24. doi:10.1016/j.cell.2018.12.015
21. Reese MG, Eeckman FH, Kulp D, Haussler D. Improved splice site detection in Genie. *J Comput Biol.* 1997;4(3):311-323. doi:10.1089/cmb.1997.4.311
22. Yeo G, Burge CB. Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals. *J Comput Biol.* 2004;11(2-3):377-394. doi:10.1089/1066527041410418
23. Martinez J, Patkaniowska A, Urlaub H, Lührmann R, Tuschl T. Single-Stranded Antisense siRNAs Guide Target RNA Cleavage in RNAi. *Cell.* 2002;110(5):563-574. doi:10.1016/S0092-8674(02)00908-X
24. Deciphering Developmental Disorders Study. Prevalence and architecture of de novo mutations in developmental disorders. *Nature.* 2017;542(7642):433-438. doi:10.1038/nature21062
25. Gurovich Y, Hanani Y, Bar O, et al. Identifying facial phenotypes of genetic disorders using deep learning. *Nat Med.* 2019;25(1):60-64. doi:10.1038/s41591-018-0279-0
26. Nakanishi K, Ascano M, Gogakos T, et al. Eukaryote-Specific Insertion Elements Control Human ARGONAUTE Slicer Activity. *Cell Reports.* 2013;3(6):1893-1900. doi:10.1016/j.celrep.2013.06.010
27. Pieper U, Webb BM, Dong GQ, et al. ModBase, a database of annotated comparative protein structure models and associated resources. *Nucleic Acids Research.* 2014;42(D1):D336-D346. doi:10.1093/nar/gkt1144
28. Miyoshi T, Ito K, Murakami R, Uchiumi T. Structural basis for the recognition of guide RNA and target DNA heteroduplex by Argonaute. *Nature Communications.* 2016;7(1). doi:10.1038/ncomms11846

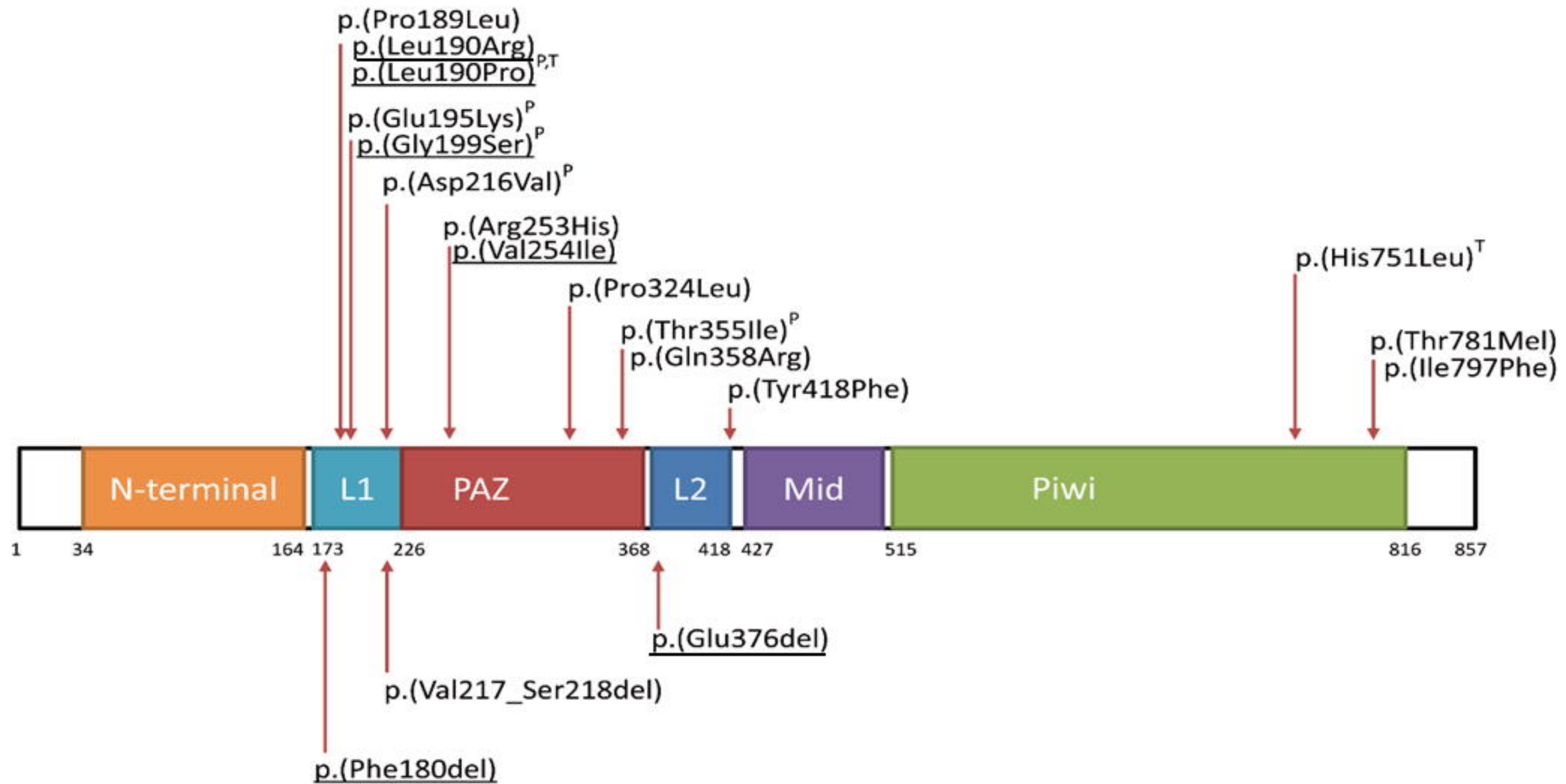
29. Kocher J-PA, Quest DJ, Duffy P, et al. The Biological Reference Repository (BioR): a rapid and flexible system for genomics annotation. *Bioinformatics*. 2014;30(13):1920-1922. doi:10.1093/bioinformatics/btu137
30. Liu X, Jian X, Boerwinkle E. dbNSFP: A lightweight database of human nonsynonymous SNPs and their functional predictions. *Human Mutation*. 2011;32(8):894-899. doi:10.1002/humu.21517
31. Schymkowitz J, Borg J, Stricher F, Nys R, Rousseau F, Serrano L. The FoldX web server: an online force field. *Nucleic Acids Research*. 2005;33(Web Server):W382-W388. doi:10.1093/nar/gki387
32. Phillips JC, Braun R, Wang W, et al. Scalable molecular dynamics with NAMD. *J Comput Chem*. 2005;26(16):1781-1802. doi:10.1002/jcc.20289
33. Karamzadeh R, Karimi-Jafari MH, Sharifi-Zarchi A, Chitsaz H, Salekdeh GH, Moosavi-Movahedi AA. Machine Learning and Network Analysis of Molecular Dynamics Trajectories Reveal Two Chains of Red/Ox-specific Residue Interactions in Human Protein Disulfide Isomerase. *Scientific Reports*. 2017;7(1). doi:10.1038/s41598-017-03966-5
34. Rashin AA, Domagalski MJ, Zimmermann MT, Minor W, Chruszcz M, Jernigan RL. Factors correlating with significant differences between X-ray structures of myoglobin. *Acta Crystallogr D Biol Crystallogr*. 2014;70(Pt 2):481-491. doi:10.1107/S1399004713028812
35. Nick Pace C, Martin Scholtz J. A Helix Propensity Scale Based on Experimental Studies of Peptides and Proteins. *Biophysical Journal*. 1998;75(1):422-427. doi:10.1016/S0006-3495(98)77529-0
36. Singh A, Manjunath LE, Kundu P, et al. Let-7a-regulated translational readthrough of mammalian AGO1 generates a microRNA pathway inhibitor. *EMBO J*. 2019;38(16). doi:10.15252/embj.2018100727
37. Parisi C, Giorgi C, Batassa EM, et al. Ago1 and Ago2 differentially affect cell proliferation, motility and apoptosis when overexpressed in SH-SY5Y neuroblastoma cells. *FEBS Lett*. 2011;585(19):2965-2971. doi:10.1016/j.febslet.2011.08.003
38. De Franco E, Watson RA, Weninger WJ, et al. A Specific CNOT1 Mutation Results in a Novel Syndrome of Pancreatic Agenesis and Holoprosencephaly through Impaired Pancreatic and Neurological Development. *Am J Hum Genet*. 2019;104(5):985-989. doi:10.1016/j.ajhg.2019.03.018
39. Uehara T, Takenouchi T, Yamaguchi Y, et al. CNOT2 as the critical gene for phenotypes of 12q15 microdeletion syndrome. *Am J Med Genet A*. 2019;179(4):659-662. doi:10.1002/ajmg.a.61068
40. Martin R, Splitt M, Genevieve D, et al. De novo variants in CNOT3 cause a variable neurodevelopmental disorder. *Eur J Hum Genet*. 2019;27(11):1677-1682. doi:10.1038/s41431-019-0413-6
41. Balak C, Benard M, Schaefer E, et al. Rare De Novo Missense Variants in RNA Helicase DDX6 Cause Intellectual Disability and Dysmorphic Features and Lead to P-Body Defects and RNA Dysregulation. *Am J Hum Genet*. 2019;105(3):509-525. doi:10.1016/j.ajhg.2019.07.010
42. Weil D, Piton A, Lessel D, Standart N. Mutations in genes encoding regulators of mRNA decapping and translation initiation: links to intellectual disability. *Biochem Soc Trans*. 2020;48(3):1199-1211. doi:10.1042/BST20200109

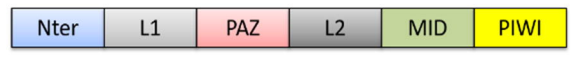
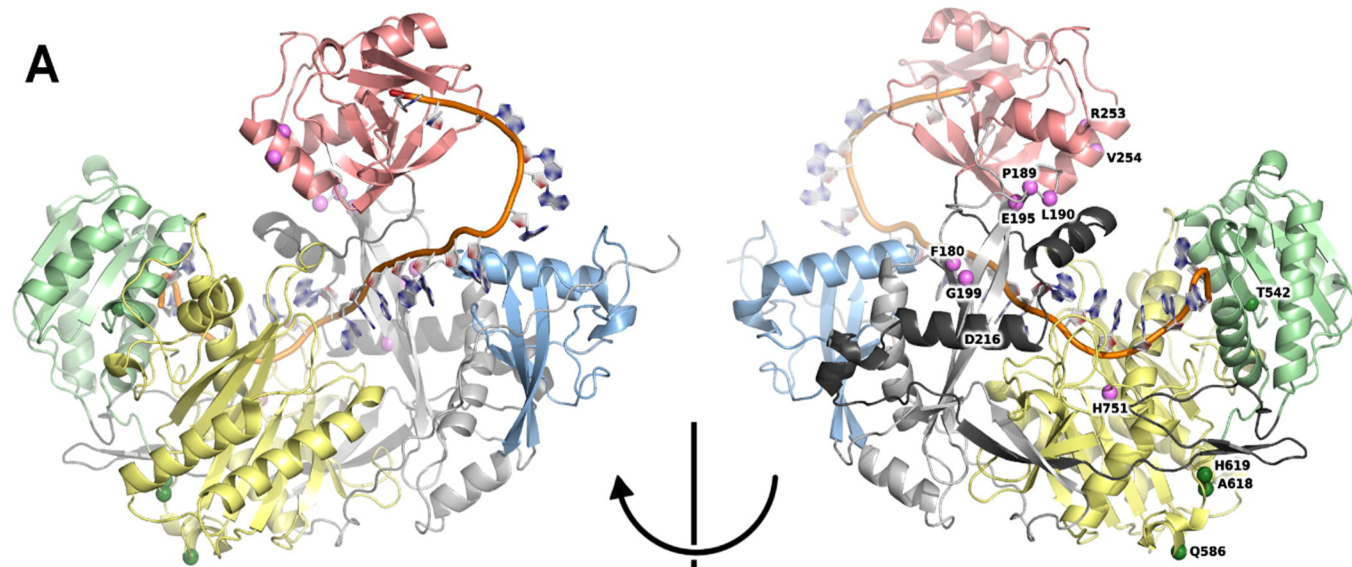
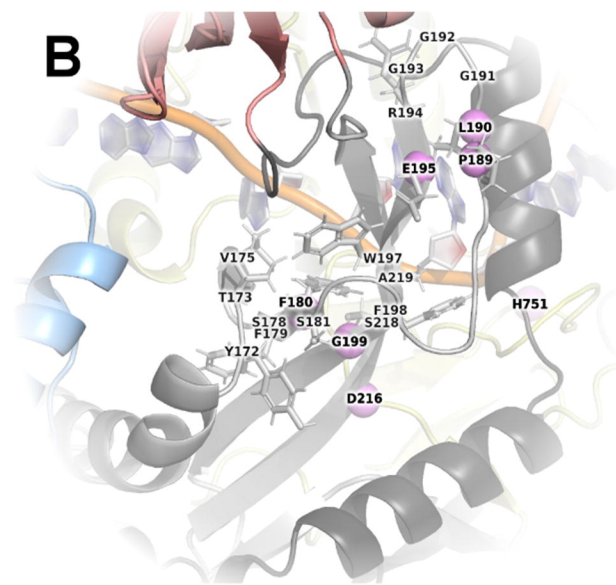
Table 1

Family	Family previously reported	Genomic position (GRCh37) - chr 1	cDNA	Protein	SIFT	PPH2 (HumVar)	CADD	Domain	$\Delta\Delta G_{fold}^{\dagger}$	Context [‡]	Dynamics Change ^α
1, 2, 3, 4, 5	-	g.36359301_36359303del	c.539_541del	p.(Phe180del)	NA	NA	NA	L1 linker	NA	Core	NA
6	-	g.36359328C>T	c.566C>T	p.(Pro189Leu)	Deleterious (score: 0.02)	Possibly Damaging (score: 0.605)	28.3	L1 linker	1.0±0.4	PAZ	PC1 -0.4, PC2 -1.6
7, 8	-	g.36359331T>G	c.569T>G	p.(Leu190Arg)	Deleterious (score: 0)	Probably Damaging (score: 1.000)	29.2	L1 linker	9.5±0.6	PAZ	NA
9	Rauch <i>et al.</i> ³	g.36359331T>C	c.569T>C	p.(Leu190Pro)	Deleterious (score: 0)	Probably Damaging (score: 1.000)	29.3	L1 linker	5.6±0.4	PAZ	PC1 -0.3, PC2 -1.1
-	Martinez <i>et al.</i> ²³	g.36359345G>A	c.583G>A	p.(Glu195Lys)	Deleterious (score: 0)	Probably Damaging (score: 0.999)	30	L1 linker	2.5±0.3	PAZ	PC1 0.1, PC2 -0.7
10, 11, 12 (Hamdan <i>et al.</i>) ⁵ , 13, 14	Sakaguchi <i>et al.</i> ⁶	g.36359357G>A	c.595G>A	p.(Gly199Ser)	Deleterious (score: 0)	Probably Damaging (score: 1.000)	29.6	L1 linker	17.9±0.7	Near Core	PC1 0.6, PC2 -0.9
-	DDD Study ²⁴	g.36359409A>T	c.647A>T	p.(Asp216Val)	Deleterious (score: 0)	Probably Damaging (score: 0.996)	32	L1 linker	2.5±0.7	Bridge R712	PC1 -0.4, PC2 -0.6
15	-	g.36359636A>G	c.650-2A>G (r.650_655del)	p.Val217_Ser218del	NA	NA	NA	L1 linker	NA	Linker and RNA H-bond	NA
16	-	g.36359746G>A	c.758G>A	p.(Arg253His)	Deleterious (score: 0.03)	Benign (score: 0.173)	25.6	PAZ	0.6±0.6	Helix Cap	PC1 0.2, PC2 -1.5
17, 18	-	g.36359748G>A	c.760G>A	p.(Val254Ile)	Deleterious (score: 0)	Benign (score: 0.092)	22.9	PAZ	-1.0±0.0	Helix Center	PC1 -0.7, PC2 -1.5
19	-	g.36360821C>T	c.971C>T	p.(Pro324Leu)	Deleterious (score: 0)	Probably Damaging (score: 1.000)	31	PAZ	2.1±0.1	PAZ	NA
-	Sanders <i>et al.</i> ⁴	g.36367118C>T	c.1064C>T	p.(Thr355Ile)	Deleterious (score: 0)	Probably Damaging (score: 0.973)	24.8	PAZ	0.4±0.6	Helix Cap	NA
20	-	g.36367127A>G	c.1073A>G	p.(Gln358Arg)	Deleterious (score: 0)	Probably Damaging (score: 0.996)	26.8	PAZ	0.3±0.2	Linker	NA
21, 22	-	g.36367180_36367182del	c.1121_1123del	p.(Glu376del)	NA	NA	NA	L2 linker	NA	Linker	NA
23	-	g.36367661A>T	c.1253A>T	p.(Tyr418Phe)	Deleterious (score: 0)	Benign (score: 0.053)	24.4	L2 linker	0.3±0.1	Linker	PC1 0.7, PC2 -0.9
24	-	g.36384011A>T	c.2252A>T	p.(His751Leu)	Deleterious (score: 0)	Probably Damaging (score: 0.982)	29.1	PIWI	-1.1±0.8	RNA H-bond	PC1 0.6, PC2 -1.2
25	-	g.36384732C>T	c.2342C>T	p.(Thr781Met)	Deleterious (score: 0)	Probably Damaging (score: 1.000)	28	PIWI	1.8±0.9	PIWI core	NA
26	-	g.36384779A>T	c.2389A>T	p.(Ile797Phe)	Deleterious (score: 0)	Probably Damaging (score: 0.969)	28.3	PIWI	3.1±0.5	PIWI near RNA	PC1 1, PC2 -0.8

Table 2

	<i>AGO1</i>		<i>AGO2</i>
	total (n=33)	%	%
ID/DD	31/31	100%	100%
Hypotonia	13/18	72%	57%
Seizures	13/28	46%	44%
Motor delay	28/30	93%	100%
Speech impairment	30/30	100%	100%
Autistic behaviour	24/30	80%	56%
Attention deficit /Hyperactivity	15/22	68%	53%
Anxiety	7/8	87.5 %	NA
Aggressiveness	11/24	78.5 %	24%
Sleeping disturbance	17/22	77%	NA
Cerebral MRI abnormalities	11/24	46%	56%
Neonatal feeding difficulties	10/23	44%	63%
Gastroesophageal reflux	3/23	13%	37%
Skeletal anomalies	5/21	23%	47%
Heart anomalies	NA	NA	33%
Strabism	4/26	15%	33%
Visual impairment	NA	NA	29%
Abnormal respiration	NA	NA	26%
Dental anomalies	NA	NA	47%



A**B****C**